The psychometric properties of the Icelandic version of the Edinburgh Postnatal Depression Scale (EPDS) when used prenatal

Linda B Lydsdottir , Louise M Howard , Halldora Olafsdottir , Marga Thome , Petur Tyrfingsson Cand , Jon Fridrik Sigurdsson

PII: S0266-6138(18)30301-2
DOI: https://doi.org/10.1016/j.midw.2018.10.009
Reference: YMIDW 2360

To appear in: *Midwifery*

Received date: 7 November 2017
Revised date: 6 September 2018
Accepted date: 16 October 2018

Please cite this article as: Linda B Lydsdottir , Louise M Howard , Halldora Olafsdottir , Marga Thome , Petur Tyrfingsson Cand , Jon Fridrik Sigurdsson , The psychometric properties of the Icelandic version of the Edinburgh Postnatal Depression Scale (EPDS) when used prenatal, *Midwifery* (2018), doi: https://doi.org/10.1016/j.midw.2018.10.009

The psychometric properties of the Icelandic version of the Edinburgh Postnatal Depression Scale (EPDS) when used prenatal

Linda B Lydsdottir MSc[1,2,*], Louise M Howard PhD[3,4], Halldora Olafsdottir MD[1,5], Marga Thome PhD[6], Petur Tyrfingsson Cand Psych[5], Jon Fridrik Sigurdsson PhD[1,2,5]

[1]*Faculty of Medicine, University of Iceland, Reykjavik, Iceland*

[2]*School of Business, Reykjavik University, Reykjavik, Iceland*

[3]*Section of Women's Mental Health, King's College London*

[4]*Institute of Psychiatry, London, United Kingdom*

[5]*Mental Health Services, Landspitali – The National University Hospital of Iceland, Reykjavik, Iceland*

[6]*Faculty of Nursing, University of Iceland, Reykjavik, Iceland*

* Corresponding author: School of Business, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland. Tel: +354-8631048. Email: linda@virk.is

Conflict of Interest:
Louise Howard is supported by the National Institute for Health Research South London and Maudsley NHS Foundation Trust specialist Biomedical Research Centre for Mental Health. All other authors declare that they have no conflicts of interest.

Ethical Approval:
Approval for the study was received from the Icelandic National Bioethics Committee (ref no. 05-107-S1) and the Icelandic Data Protection Authority (ref no. S2589).

Funding Sources:
This work was supported by the Icelandic Centre for Research (RANNIS) 50427021, 50427022, 50427023, 100740001; The University of Iceland Research Fund; The Landspitali–University Hospital Research Fund and the Wyeth Research Fund. These grants had no further role in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication.

Clinical Trial Registry and Registratin number:
Not applicable.

Objective: To evaluate the psychometric properties of the Icelandic version of the Edinburgh Postnatal Depression Scale (EPDS) when used prenatal, explore the dimensionality of the scale and describe its effectiveness in identifying depression.

Design: A sample of Icelandic women filled in the EPDS at week 16 gestation, week 24 and week 36. If screened positive in week 16 they were asked to attend a psychiatric diagnostic interview 2-4 weeks later. Every10th woman screened negative was also asked to attend an interview.

Setting: Antenatal clinics at primary health care centres in Iceland.

Participants: In total, 2512 women receiving prenatal care participated in the study. At week 16 gestation, 2397 women filled in the Edinburgh Postnatal Depression Scale, 2025 at week 24, and 1756 at week 36. 474 women attended diagnostic interview two to four weeks after screening.

Measurements and findings: Internal reliability, convergent validity and test-retest correlation of the Icelandic version of the Edinburgh Postnatal Depression Scale appeared acceptable. An exploratory factor analysis supported a one-factor structure of the Edinburgh Postnatal Depression Scalethat was confirmed by confirmatory factor analysis showing best fit for one general factor with two group factors. A cut-off score of 11 or higher had specificity of .89, sensitivity of .80 and positive predictive value of .44.

Key conclusions and implications for practice: The Icelandic version of the Edinburgh Postnatal Depression Scale is a valid and reliable one-dimensional instrument suitable to screen for depression prenatally. We recommend using score 11 or higher as a cut-off. If women score 11, they should be re-assessed two weeks later, but if they score 12 or higher, they should be referred directly for a further assessment. A time gap of two to four weeks does weaken the scale's ability to discriminate between those suffering from Major Depression and those who screen negative.

Keywords: EPDS, prenatal care, depression, factor analysis, statistical, psychometric properties

Abbreviations: EPDS, EFA, CFA, DASS, MINI+

Introduction

Health care professionals are mostly aware of the burden of depressive symptoms many women experience in prenatal care. Experiencing depressive symptoms may not only have serious consequences for the mother, but also for the unborn baby (Stein et al., 2014). The Edinburgh Postnatal Depression Scale (EPDS; (Cox et al., 1987) is the most widely studied screening instrument for depressive symptoms perinatally (Gibson et al., 2009). It is a ten-item self-report questionnaire without somatic items found in other known screening instruments, which could be misleading as they may be pregnancy related. The EPDS was originally developed for screening of postpartum depression (Cox et al., 1987), but soon after it was developed, it was used to screen for depression prenatally as well (Murray and Cox, 1990). Today it is the most studied screening instrument for depression both pre- and postnatal, has been translated and validated in numerous countries and is used worldwide in research and practice (Cox and Holden, 2003; Eberhard-Gran et al., 2001; Gibson et al., 2009; Kozinszky & Dudas, 2015; Matthey & Agostini, 2017). Systematic reviews of its psychometric properties have shown that it is not advisable to use universal cut-off scores because of cultural differences (Eberhard-Gran et al., 2001; Gibson et al., 2009; Kozinszky & Dudas, 2015), indicating the importance of carrying out a validity evaluation in each country before deciding to use the scale, as a screening tool, in a clinical setting or in research. The Icelandic version has not been validated for prenatal use, but in spite of that, several antenatal clinics already use the scale for prenatal screening and in recently released guidelines for the Primary Health Care in Iceland (2017), screening with the EPDS for depression is recommended during the first visit (week 16).

Although Cox et al. (1987) originally described EPDS as unidimensional, several studies have looked at its factor structure, reporting either two factors (depression and anxiety) or three factors (depression, anxiety and anhedonia), but with large variation in item-factor structures. This refers both to the postnatal (Astbury et al., 1994; Berle et al., 2003; Matthey, 2008; Phillips et al., 2009; Pop et al., 1992; Ross et al., 2003; Swalm et al., 2010; Tuohy & McVey, 2008) and prenatal periods (Adouard et al., 2005; Agampodi & Agampodi, 2013; Bowen et al., 2008; Brouwers et al., 2001; Jomeen & Martin 2005; Jomeen & Martin 2007; Montazeri et al., 2007; Ross et al., 2003; Swalm et al., 2010; Töreki et al., 2013; Zhong et al., 2014). Many studies reporting more than one factor have advocated the use of the EPDS as a multidimensional screening instrument for both depression and anxiety (Bina & Harrington, 2016; Matthey, 2008). Conclusions should be made carefully as methodological limitations have been found in many studies, like small sample sizes and different statistical methods, e.g. some using principal component analysis (PCA) while others have used exploratory factor analysis (EFA) or confirmatory factor analysis (CFA), and the use of inappropriate estimation methods for ordered scales.

Our purpose was to evaluate if the EPDS is a suitable screening tool for prenatal depression among Icelandic women. The aim was to evaluate the psychometric properties of the scale and explore its dimensionality for three different measuring points during pregnancy, controlling for some of the limitations found in other studies. In the paper, we also describe the scale's effectiveness in identifying depression prenatal.

Methods

Participants

The participants were women attending antenatal clinics at primary health care centres in Iceland between September 2006 and May 2011. The size of the Icelandic population was 353.070 thousand individuals the 20th July 2018 and in 2017, 4.071 children were born, 2.112 males and 1.959 females (Statistics Iceland, 22. August 2018). Antenatal clinics are free of charge and the attendance rate is about 100%. Fifteen antenatal clinics are located in the capital region (the greater Reykjavik area) and ten of them participated in the study. One clinic located in Iceland's second largest urban area, Akureyri, also participated. Inclusion criteria were (a) being pregnant, (b) being at least 16 years of age, and (c) being able to read and speak Icelandic fluently. Women with history of schizophrenia or other psychotic illnesses and women with significantly impaired cognitive functioning were excluded from the study. If a participating woman experienced spontaneous abortion during pregnancy she was excluded from the study. In total 2512 pregnant women participated in the study (mean age being 29 (17-47), SD = 5.16).

Instruments

The Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987) is a 10-item self-rating scale designed to screen for postpartum depression. The scale covers the most common symptoms of depression, without somatic symptoms such as fatigue and change in appetite, which are common symptoms in women pre- and postnatally. Scoring for each item is from 0 to 3, high scores indicating more symptoms of depression (ranging from 0 to 30). The EPDS has been validated in many countries and has good psychometric properties when used both pre- and postnatal (Gibson et al., 2009; Murray and Cox, 1990). The EPDS was originally translated from English to Icelandic by one of the authors of this study (MT). For this study two experienced clinicians, a psychologist and a psychiatrist, re-translated the scale. It was then back-translated to English by a professor in Clinical Psychology. When there was a disagreement between the back-translated item and the original item, a third psychologist was asked to translate the item and then it was back-translated again. Pretesting was carried out on postpartum women. A cut-off score of 15 or more on the EPDS has been recommended when reporting probable major depression in antenatal English-speaking women (Matthey et

al., 2006). For this study a lower cut-off score was chosen, i.e. 12 or more, after a thorough review of validation studies on the EPDS (Gibson et al., 2009; Kozinszky & Dudas, 2015).

The Depression Anxiety and Stress Scales (DASS; Lovibond and Lovibond, 1995a) is a set of three self-report scales used to screen for symptoms of depression, anxiety and stress, each consisting of 14 items. The DASS has been shown to be reliable (alpha coefficients of 0.91, 0.81 and 0.89, for the Depression, Anxiety and Stress scales, respectively) with good validity (Lovibond and Lovibond, 1995b). The reliability of the Icelandic version is good (0.90, 0.86 and 0.91, for the Depression, Anxiety and Stress scales, respectively; Gudjonsson et al., 2009). The following cut-off scores were used in this study: ≥10 for depression and ≥8 for anxiety, which indicate a mild state of either condition (Lovibond and Lovibond, 1995a). The stress scale was not used in this study.

The Mini-International Neuropsychiatric Interview plus (MINI+; Lecrubier et al., 1997) is a standard diagnostic interview that contains 26 modules for the major axis I psychiatric disorders in DSM-IV TR and ICD-10. Administration takes approximately 20 minutes by experienced clinicians. The MINI+ has acceptable test-retest and inter-rater reliability and has been validated against the SCID for the DSM-III R and the CIDI for the ICD-10 (Lecrubier, et al., 1997; Sheehan, Lecrubier et al., 1997; Sigurðsson, 2008).

A questionnaire about background variables was designed for this study, including questions about age, marital status, education, working status and number of children.

Procedure

This study is part of a longitudinal project on pregnant women attending prenatal care in Iceland. At week 12 to 14 gestation, women were asked to participate and all those who agreed were asked to give signed informed consent. They were then screened for depression and anxiety three times during pregnancy (around week 16, 25 and 36) using the EPDS and the DASS. Midwives and nurses in prenatal care carried out the recruitment and screening under the supervision of one of the main researchers (LBL).

The women who were screened with depressive and/or anxiety symptoms were contacted by the researchers two to four weeks after screening and asked to attend a psychiatric diagnostic interview. Every 10th woman from each clinic who screened negative was also contacted and asked to attend a psychiatric diagnostic interview. Experienced female

clinicians (psychologist, midwife and psychiatrist) who conducted the interviews, using the MINI+ were all blind to the participants' EPDS and DASS scores. The inter-rater reliability between interviewers on the MINI+ was Kappa = .86 ($p < 0.001$), 95% CI (.75, .97). According to Landis and Koch (1977) Kappa values from 0.40 to 0.59 are considered moderate, 0.60 to 0.79 substantial and 0.80 and above outstanding. To see if the results differed depending on whether there was a delay between the test administration and the interview, the last 102 women interviewed were asked to answer the EPDS immediately after being interviewed.

Statistical analysis

The Statistical Package for Social Sciences (SPSS) version 21 (IBM Corp, 2012) was used for all statistical analysis except for the factor analysis. A descriptive analysis was carried out on the socio-demographic characteristics of the women attending psychiatric diagnostic interviews. The sample was then categorised into clinical and nonclinical sample based on a MINI+ diagnosis of Major Depression. To address the discriminant validity of the EPDS and the DASS, independent t-tests were calculated on the mean scores between the clinical and the non-clinical samples. Hierarchical multiple regression was carried out to test the convergent validity of the EPDS. The internal reliability of the EPDS was calculated using Cronbach's alpha.

To analyse the factor structure of the EPDS, an Exploratory Factor Analysis (EFA) was first performed on a randomly selected subsample. Some studies using EFA have not taken into account the fact that they are handling ordinal variables, thus treating them as interval scales. This may affect the outcomes, as it may underestimate the strength of the relationships between the variables and therefore give rise to spurious multidimensionality (Baglin, 2014). Therefore, the polychoric correlation technique was chosen, as it may correct for this bias (Timmerman & Lorenzo – Seva, 2011). We used FACTOR 9.2 (Ferrando & Lorenzo-Seva, 2017; Lorenzo-Seva & Ferrando, 2006) conducting parallel analysis based on Minimum Rank Factor Analysis (MRFA; Shapiro & Ten Berge, 2002) as recommended by Timmerman and Lorenzo-Seva (2011). Parallel analysis has been found to be superior to other methods when identifying the number of factors (Baglin, 2014). The conventional method, the Kaiser criteria (eigenvalue <1), has been shown to overestimate the number of factors in the data (Ruscio and Roche 2012). Oblique rotation was then used and loadings below 0.4 are not reported.

Confirmatory factor analysis (CFA) was then performed using the R 3.0.2 Lavaan (Rosseel, 2012) to test the factor structure of the EPDS. The data was analysed using the Weighted Least Squares Mean and Variance (WLSMV)

robust estimator since the data was ordinal. Chi-square statistic ($\chi 2$) was used to test the fit of each model, although it is problematic in large samples (Jöreskog, 1969), making it difficult to retain the null hypothesis. Therefore, the following statistics were also used to test the models' fit: the comparative fit index (CFI; Bentler, 1990) and the Tucker-Lewis index (TLI; Tucker & Lewis, 1973), with values above 0.95 indicating a good fit (Hu and Bentler, 1999), and the Root Mean Square Error of Approximation (RMSEA), with value under 0.05 indicating adequate fit (Schumacker & Lomax, 1996). Further, the Weighted Root Mean Square Residual (WRMR) was used, as it is highly appropriate for data that is not normally distributed (Muthén & Muthén, 1998-2015), with the lowest value indicating the most adequate fit (Cook et al., 2009). To test for changes in the validity of the EPDS depending on different time periods in pregnancy, the CFA was performed using data from gestation weeks 16, 25 and 36. The following five models were tested: (1) one-factor model including all 10 items (Cox et al., 1987); (2) Jomeen and Martin (2005) two-factor model with items 1, 2, 6, 7, 8 and 9 loading on one factor (depression) and items 3, 4, and 5 loading on a second one (anxiety); (3) Phillips et al. (2009) two-factor model with items 1, 2, 6, 7, 8, 9 and 10 loading on one factor (depression) and items 3, 4, 5 on a second one (anxiety); (4) Zhong et al. (2014) two-factor model with items 1 and 2 loading on one factor (anhedonia) and items 3, 4, 5, 6, 7, 8, 9, 10 loading on a second factor (anxiety and depression); (5) Reichenheim et al., (2011) three-factor model with items 1, 2, and 6 loading on one factor (anhedonia), items 3, 4, and 5 loading on the second factor (anxiety) and items 7, 8, 9 and 10 loading on the third factor (depression). Finally, we tested several bi-factor models (Gibbons & Hedeker, 1992). We tested four bi-factor models based on the Jomeen and Martin (2005) two-factor model, the Phillips et al. (2009) two-factor model, the Zhong et al. (2014) two-factor model, and the Reichenheim et al. (2011) three-factor model.

Finally, on the basis of the factor analysis, a Receiver Operating Characteristics Analysis (ROC) was carried out with regards to a diagnosis on the MINI+, and the sensitivity, specificity, likelihood ratio and the positive predictive value were calculated

Results

Description of the sample

In total, 2512 women participated in the study. The numbers of women filling in scales at week 16 gestation, week 24 and week 36, and the percentage of screened positive women, are shown in Table 1. Of 396 women who screened positive in week 16 on the EPDS, the DASS-Depression Scale or the DASS-Anxiety Scale, 273 (69% response rate) attended a psychiatric diagnostic interview two to four weeks later. Of 324 women asked to participate in the screen

negative group, 201 (62% response rate) attended a psychiatric interview. Thus, in total 474 women attended psychiatric interview two to four weeks after having answered screening questionnaires in week 16. Of them 189 (40%) women were pregnant with their first child, 429 (90.5%) were married or cohabiting, 240 (50.6%) had a university degree, 307 (64.8%) were working and 41 (8.6%) were without work because of sickness or unemployment.

-Table 1 about here-

Reliability

The internal reliability coefficients for the EPDS were good (Cronbach & Shavelson 2004), alpha being 0.85 (95% CI 0.84 - 0.86), 0.84 (95% CI 0.83 – 0.85) and 0.84 (95% CI 0.83 – 0.85) at weeks 16, 24 and 36 gestation, respectively. In total, 102 women answered the second administration of the EPDS two to four weeks later. Test-retest correlation was acceptable, 0.68.

Validity

The EPDS mean for the total sample screened in week 16 was 5.5 (SD=4.1). To assess the scale's discriminant validity, a t-test was carried out for those diagnosed with Major Depression (M=12.9, SD = 4.8) and the screen negative group (M=7.9, SD=4.8). The difference between the two groups was large and significant, $t(473)=9.01$, $p<.001$, indicating good discriminative validity.

To establish Convergent validity of the EPDS a hierarchical multiple regression was carried out with scores on the EPDS as the outcome variable and scores from the three DASS scales as predictor variables. The results are presented in Table 2. The Table shows that the DASS scales explained 66% of the variance of the EPDS, of which the DASS Depression scale alone explained 59%, indicating that the scale covers almost all the variation and that the DASS Anxiety and DASS Stress scales add little to the explained variance.

-Table 2 about here-

Explorative factor analysis

Explorative factor analyses were conducted on the EPDS for weeks 16, 25, and 36 gestation in a randomly selected subsample. Mardia's test showed no significant skewness ($p = 1$), but evidence for excessive kurtosis was found on all measurements points, $p < .001$ (246.70, 282.25 and 312.25 for weeks 16, 25 and 36 gestation, respectively); supporting the decision to use polychoric correlations. The appropriateness was verified by the significant Bartlett's test for the three time points ($\chi2=8207.2$; $\chi2=6780.7$; $\chi2=5776.8$ ($df = 45$), $p<.001$). The Kaier-Mayer-Olkin measures for weeks 16, 25, and 36 were .906, .902 and .890, respectively. Thus, both methods indicating that the database is suitable for exploratory factor analysis. Eigenvalues above 1 indicated a two-factor solution, but the mean of variance extracted, the 95th percentile of random percentage of variance and the Scree plot, indicated a one-factor solution. This one factor accounts for approximately 81.84%, 78.44% and 78.26% of the variance. The RMSR was .055, .060 and .063 for the three time points, respectively. As Table 3 describes, all the items loaded above .5 on all the measurement points.

-Table 3 about here-

To test the dimensionality of the EPDS with confirmatory factor analysis, several models from other studies were tested, as shown in Table 4. First, a one-factor model was tested in accordance with our EFA, secondly, the Jomeen and Martin (2005) two-factor model, followed by the Phillips et al. (2009) two-factor model, the Zhong et al. (2014) two-factor model and the Reichenheim et al. (2011) three-factor model. Correlations between factors were allowed for models with two or more factors. Finally, we tested Jomeen and Martin (2005), Phillips et al. (2009), Zhong et al. (2014) and Reichenheim et al. (2011), as bi-factor models. All models yielded large significant Chi-square statistics, probably due to the large sample size.

-Table 4 about here-

All models showed adequate fit to the data with the CFI, TLI and RMSEA statistics being acceptable, as can be seen in Table 4. In all cases the bi-factor models showed more optimal fit to the data indicating a general factor. A bi-factor version of the Zhong et al. (2014) model showed the best fit in weeks 16, 25 and 36.

ROC analyses

A ROC analysis revealed that the EPDS discriminates reasonably well between pregnant women diagnosed with depression and those not diagnosed, as can be seen in Figure 1 (area under the curve (auc) =.947, 95% CI =.897-.964), when there was no delay between test administration and the diagnostic interview (n=102).

-Figure 1 about here-

The Youden Index (J; Youden, 1950) indicated that the optimal cut-off value would be 9 or more with 100 in sensitivity and 76.09 in specificity as can be seen in Table 5, which also shows the likelihood ratio and the positive predictive values of the optimal cut-off values. The positive predicted value is only .31 when using cut-off 9, indicating a low probability of depression among women with a positive test, and the likelihood ratio is around 4, indicating a small increase in the likelihood of depression. Thus, the score of 9 or higher as recommended cut-off is not acceptable. A cut-off score of 11 gives us acceptable sensitivity, .80, with higher specificity, .89, positive predicted values, .44, and likelihood ratio, 7.36, which indicates moderate likelihood of depression. With higher cut-off at 12 we have a likelihood ratio indicating conclusive increase in the likelihood of Major Depression, with high specificity, .93, positive predictive value, .54, and sensitivity, .70.

-Table 5 about here-

When 2 to 4 weeks passed between test administration of the EPDS and the diagnostic interview, it appeared that the EPDS discriminated less well between those diagnosed with Major Depression and the screen negative group, as can be seen in Figure 2 (area under the curve (auc) =.809, 95% CI =.771-.843). The Youden Index indicates that the optimal cut-off value is 10 or more, higher than when no time passed between administration and the interview, but the specificity, positive predicted values and likelihood ratio are lower, as can be seen in Table 5.

-Figure 2 about here-

Discussion

In this study the Icelandic version of the EPDS was validated in a sample of pregnant women and its dimensionality explored. The results indicate good reliability and validity of the scale. The results show that internal consistency is good, which is in line with other studies on the EPDS (Gibson et al., 2009) and test-retest correlation is acceptable.

The means and standard deviations do not show any deviations from what has been found in other studies, both for clinical and non-clinical samples (Gibson et al., 2009; Kozinszky & Dudas, 2015) and good convergent validity was reported. When studying construct validity, our results from the exploratory factor analysis on the EPDS showed one-factor structure. We used a polychoric correlation technique, taking into account that the EPDS has an ordinal scale. Treating ordinal data as interval data, as older studies have done, may be appropriate if there are at least five scale items and preferably seven, but not when the instruments consist of fewer categories, as in the EPDS. In such cases, one-factor measurement models are too often rejected, resulting in a bias such as spurious multidimensionality (Baglin, 2014).

Although our EFA using polychoric correlation revealed a one-factor solution, we included several types of two-factor models in our CFA. That analysis did not support a one-factor model, but supported various types of two-factor models, showing little difference between them on several fit indexes. This problem has been reported when testing dimensionality of other psychological measures (Reise et al., 2007). Most studies do find some evidence of multidimensionality when using CFA even after reporting one factor with a large first eigenvalue (Chen et al., 2006). A likely explanation for why this discrepancy so often appears in the literature is that a construct like depression is very complex, involving diverse indicators. Therefore, some studies may report evidence for a single structure while others may uncover evidence of multidimensionality, making the debate on-going (Reise et al., 2007). We therefore tested several bi-factor models, with one showing the best fit to our data, one general factor with two group factors (item 1 and 2 vs items 3-10). Only minor differences were found between the three measuring points, at weeks 16, 24, and 36 gestation, indicating a clear one factor structure in the prenatal period.

When analysing how well the EPDS discriminates between Icelandic women diagnosed with Major Depression and those not diagnosed, we found that it did reasonably well when there was no delay between screening and the diagnostic interview. Bergink et al. (2011) recommended using a score of 11 or higher as cut-off in a general sample. If we choose the same cut-off in our sample we would have acceptable sensitivity, .80, with a higher specificity, .89, positive predicted value 44.4 and likelihood ratio of 7.36, which indicated moderate likelihood of depression. With a higher cut-off score of 12 or higher, we got a positive predicted value 53.8 and likelihood ratio of 10.37, indicating a conclusive increase in the likelihood of Major Depression, with high specificity of .93 but lower sensitivity of .70, thus leaving a larger group of depressed women undiagnosed. A low positive predictive value is of concern, as it indicates how likely it is that women screened positive are indeed depressive. Other studies have also reported a low

positive predictive value for the EPDS, usually around 50-60% or lower (Kozinszky and Dudas 2015; Matthey et al., 2006; Matthey & Agostini, 2017). When using the EPDS to screen women in the prenatal period, professionals therefore need to be aware that screen-positive women may not necessarily be suffering from Major Depression as has been reported elsewhere (Lydsdottir et al., 2014), making it important to assess them further.

A clear strength of our study is the large community sample of pregnant women, but there are some limitations. Women with more severe symptoms of depression may have dropped out and in the middle of the data collection, the economic collapse hit Iceland. Studies have shown that following the collapse the stress levels and the risks of depressive symptoms increased in Icelandic women (Hauksdottir et al., 2013; McClure, 2014), thus indicating that mental health among Icelandic pregnant women may also have been affected. We conclude that the Icelandic version of the EPDS is reliable and valid when used in a sample of pregnant women. Our results indicate clearly that the list has one general factor measuring depression on all three occasions during pregnancythus we do not recommend that the list should be used to screen for anxiety. We recommend using score 11 or higher as a cut-off when screening for depression in a general sample of Icelandic women. We further suggest that women scoring 12 or higher should be referred to further assessment, but those with score of 11 should be reassessed two weeks later. A time gap of two to four weeks does weaken the scale's ability to discriminate between those suffering from Major Depression and those who screen negative. We further suggest that until more data exist from other gestational periods, the EPDS should only be used as a screening tool around week 16 prenatal.

Acknowledgements:
We thank the Primary Health Care of Reykjavik the Capital area of Iceland, the Primary Health care of Akureyri and the Perinatal Service of Landspitali–The National University Hospital of Iceland. We also thank the women who participated in the study, research assistants and others who contributed to the study and the Icelandic Centre for Research (RANNIS), the University of Iceland Research Fund, the Landspitali–University Hospital Research Fund and the Wyeth Research Fund for their funding. Finally, we want to thank Hjalti Einarsson, Pétur Ingi Pétursson, Maria Hronn Einarsdottir, Thorbjorg Sveinsdottir and Sigridur Sia Jonsdottir for assistance with the research and Anton Orn Karlsson for consultation on data analysis.

References

Adouard, F., Glangeaud-Freudenthal, N. M. C., & Golse, B., 2005. Validation of the Edinburgh postnatal depression scale (EPDS) in a sample of women with high-risk pregnancies in France. Archives of Women's Mental Health, 8(2), 89-95.

Agampodi, S. B., & Agampodi, T. C., 2013. Antenatal depression in Anuradhapura, Sri Lanka and the factor structure of the Sinhalese version of Edinburgh post-partum depression scale among pregnant women. PloS one, 8(7), e69708.

Astbury, J., Brown, S., Lumley, J., & Small, R., 1994. Birth events, birth experiences and social differences in postnatal depression. Australian and New Zealand Journal of Public Health, 18, 176-184.

Baglin, J., 2014. Improving your exploratory factor analysis for ordinal data: a demonstration using FACTOR. Practical Assessment, Research & Evaluation, 19(5), 2.

Bentler, P. M., 1990. Comparative fit indexes in structural models. Psychological Bulletin, 107(2), 238.

Bergink, V., Kooistra, L., Lambregtse-van den Berg, M.P., Wijnen, H., Bunevicius, R., van Baar, A. & Pop V., 2011. Validation of the Edinburgh Depression Scale during pregnancy. Journal of Psychosomatic Research, 70(4), 385-389.

Berle, J. Ø., Aarre, T. F., Mykletun, A., Dahl, A. A., & Holsten, F., 2003. Screening for postnatal depression: validation of the Norwegian version of the Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. Journal of Affective Disorders, 76(1), 151-156.

Bina, R., & Harrington, D., 2016. The Edinburgh Postnatal Depression Scale: screening tool for postpartum anxiety as well? Findings from a confirmatory factor analysis of the Hebrew version. Maternal and Child Health Journal, 20(4), 904-914.

Bowen, A., Bowen, R., Maslany, G., & Muhajarine, N., 2008. Anxiety in a socially high-risk sample of pregnant women in Canada. The Canadian Journal of Psychiatry, 53(7), 435-440.

Brouwers, E. P., van Baar, A. L., & Pop, V. J., 2001. Does the Edinburgh Postnatal Depression Scale measure anxiety? Journal of Psychosomatic Research, 51(5), 659-663.

Chen, F.F., West, S.G. & Sousa, K.H., 2006. A comparison of bifactor and second-order models of quality of life. Multivariate Behavioral Research, 41(2): 189-225.

Cook, K. F., Kallen, M. A., & Amtmann, D., 2009. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. Quality of Life Research, 18(4), 447-460.

Cox, J., & Holden, J., 2003. Perinatal mental health: A guide to the Edinburgh Postnatal Depression Scale (EPDS). Royal College of Psychiatrists.

Cox, J. L., Holden, J. M., & Sagovsky, R., 1987. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. The British Journal of Psychiatry, 150(6), 782-786.

Cronbach, L. J., & Shavelson, R. J., 2004. My current thoughts on coefficient alpha and successor procedures. Educational and Psychological Measurement, 64(3), 391-418.

Eberhard- Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., & Ove Samuelsen, S., 2001. Review of validation studies of the Edinburgh Postnatal Depression Scale. Acta Psychiatrica Scandinavica, 104(4), 243-249.

Ferrando, P. J., & Lorenzo-Seva, U., 2017. Program FACTOR at 10: Origins, development and future directions. Psicothema, 29, 236-240.

Gibson, J., McKenzie- McHarg, K., Shakespeare, J., Price, J., & Gray, R., 2009. A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. Acta Psychiatrica Scandinavica, 119(5), 350-364.

Gibbons, R. D., & Hedeker, D. R., 1992. Full-information item bi-factor analysis. Psychometrika, 57(3), 423-436.

Gudjonsson, G. H., Sigurdsson, J. F., Lydsdottir, L. B., & Olafsdottir, H., 2008. The relationship between adult romantic attachment and compliance. Personality and individual differences, 45, 276-280.

Hauksdóttir, A., McClure, C., Jonsson, S.T., Ólafsson, Ö., & Valdimarsdóttir, U.V., 2013. Increased Stress Among Women Following an Economic Collapse—A Prospective Cohort Study. American Journal of Epidemiology, 177(9), 979–988,

Hu, L. T., & Bentler, P. M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: a Multidisciplinary Journal, 6(1), 1-55.

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Jomeen, J. & Martin, C. R., 2005. Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. Journal of Reproductive and Infant Psychology, 23, 143-154.

Jomeen, J. & Martin, C. R., 2007. Replicability and stability of the multidimensional model of the Edinburgh Postnatal Depression Scale in late pregnancy. Journal of Psychiatric and Mental Health Nursing, 14(3), 319-324.

Jöreskog, K. G., 1969. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34(2), 183-202Kozinszky, Z. & Dudas, R. B., 2015. Validation studies of the Edinburgh Postnatal Depression Scale for the antenatal period. Journal of Affective Disorders, 176, 95-105.

Landis, J. R., & Koch, G. G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159-174.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., & Dunbar, G. C., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. European Psychiatry, 12(5), 224-231.

Lorenzo-Seva, U., & Ferrando, P. J., 2006. FACTOR: A computer program to fit the exploratory factor analysis model. Behavior Research Methods, 38(1), 88-91.

Lovibond, S. H., & Lovibond, P. F. 1995a. Manual for the Depression Anxiety Stress Scales, 2nd edition. Sydney, Psychology Foundation.

Lovibond, P. F., & Lovibond, S. H., 1995b. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy, 33, 335–343.

Lydsdottir L.B., Howard L.M., Olafsdottir H., Thome M., Tyrfingsson P., Sigurdsson J.F., 2014. The mental health characteristics of pregnant women with depressive symptoms identified by the Edinburgh Postnatal Depression Scale. Journal of Clinical Psychiatry, 75, 393–398.

McClure, C.B., 2014. Mental health and health behaviors following an economic collapse The case of Iceland. (Unpublished Doctoral dissertation). University of Iceland, Reykjavik, Iceland. Montazeri A, Torkan B, & Omidvari S., 2007. The Edinburgh Postnatal Depression Scale (EPDS): translation and validation study of the Iranian version. BMC psychiatry, 7(1), 11.

Matthey, S., 2008. Using the Edinburgh Postnatal Depression Scale to screen for anxiety disorders. Depression and Anxiety, 25(11), 926-931.

Matthey, S., & Agostini, F., 2017. Using the Edinburgh Postnatal Depression Scale for women and men—some cautionary thoughts. Archives of Women's Mental Health, 20(2), 345-354.

Matthey, S., Henshaw, C., Elliott, S., & Barnett, B., 2006. Variability in use of cut-off scores and formats on the Edinburgh Postnatal Depression Scale–implications for clinical and research practice. Archives of Women's Mental Health, 9(6), 309-315.

Murray, L., & Cox, J.L., 1990. Screening for depression during pregnancy with the Edinburgh Postnatal Depression Scale (EPDS). Journal of Reproductive and Infant Psychology 8, 99-107.

Muthén, L. K., Muthén, B. O. 1998-2015. Mplus User's Guide, 7th edition. Los Angeles, CA: Muthén & Muthén.

Phillips, J., Charles, M., Sharpe, L., & Matthey, S., 2009. Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants. Journal of Affective Disorders, 118(1), 101-112.

Pop, V.J., Komproe, I.H., & van Son, M.J., 1992. Characteristics of the Edinburgh Post Natal Depression Scale in the Netherlands. Journal of Affective Disorders, 26, 105-110.

Reichenheim, M. E., Moraes, C. L., Oliveira, A. S., & Lobato, G., 2011. Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): empirical evidence for a general factor. BMC Medical Research Methodology, 11(1), 93.

Reise, S. P., Morizot, J., & Hays, R. D., 2007. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Quality of Life Research, 16(1), 19-31.

Ross, L. E., Evans, S. G., Sellers, E. M., & Romach, M. K., 2003. Measurement issues in postpartum depression part 1: anxiety as a feature of postpartum depression. Archives of Women's Mental Health, 6(1), 51-57.

Rosseel, Y., 2012. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). Ghent, Belgium: Ghent University.

Ruscio, J., & Roche, B., 2012. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. Psychological Assessment, 24(2), 282.

Schumacker, R.E., & Lomax, R.G., 1996. A Beginner's Guide to Structural Equation Modeling. Mahwah, NJ: Lawrence Erlbaum.

Shapiro, A., & Ten Berge, J. M., 2002. Statistical inference of minimum rank factor analysis. Psychometrika, 67(1), 79-94.)

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A., & Dunbar, G. C., 1997. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. European Psychiatry, 12(5), 232-241.

Sigurdsson, B.H., 2008. Samanburdur á tveimur stodludum gedgreiningarvidtolum og tveimur sjalfsmatskvordum: MINI, CIDI, PHQ og DASS (Unpublished master's thesis). University of Iceland, Reykjavik, Iceland.

Stein, A., Pearson, R. M., Goodman, S. H., Rapa, E., Rahman, A., McCallum, M., ... & Pariante, C. M., 2014. Effects of perinatal mental disorders on the fetus and child. The Lancet, 384(9956), 1800-1819.

Swalm, D., Brooks, J., Doherty, D., Nathan, E., & Jacques, A., 2010. Using the Edinburgh postnatal depression scale to screen for perinatal anxiety. Archives of Women's Mental Health, 13(6), 515-522.

Timmerman, M. E., & Lorenzo-Seva, U., 2011. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychological Methods, 16(2), 209.

Tucker, L. R., & Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38(1), 1-10.

Tuohy, A., & McVey, C., 2008. Subscales measuring symptoms of non-specific depression, anhedonia, and anxiety in the Edinburgh Postnatal Depression Scale. British Journal of Clinical Psychology, 47(2), 153-169.

Töreki, A., Andó, B., Keresztúri, A., Sikovanyecz, J., Dudas, R. B., Janka, Z., ... & Pál, A., 2013. The Edinburgh Postnatal Depression Scale: translation and antepartum validation for a Hungarian sample. Midwifery, 29(4), 308-315.

Youden, W. J., 1950. Index for rating diagnostic tests. Cancer, 3(1), 32-35.

Zhong, Q., Gelaye, B., Rondon, M., Sánchez, S. E., García, P. J., Sánchez, E., ... & Williams, M. A. (2014). Comparative performance of patient health questionnaire-9 and Edinburgh Postnatal Depression Scale for screening antepartum depression. Journal of Affective Disorders, 162, 1-7

Table 1. Number of women answering lists at week 16 gestation, week 24 and week 36 and the percentage of screened positive women

	Week 16 (n=2397)	Week 24 (n=2025)	Week 36 (n=1756)
All lists	9.7%	8.8%	8.7%
EPDS	9.6%	9.3%	9.2%
Dass Depression	9.4%	7.7%	7.1%
Dass Anxiety	9.9%	11.9%	11.1%

EPDS: the Edinburgh Postnatal Depression Scale

Table 2. DASS Scales scores incremental explained variance of EPDS

	B	SE B	β	95% CI
Step 1				
Constant	4.94	.22		4.51-5.36
DASS Depression	0.52	.02	.76*	0.48-0.56
Step 2				
Constant	3.63	.24		3.16-4.09
DASS Depression	0.30	.03	.44*	0.25-0.36
DASS Anxiety	0.17	.03	.19*	0.10-0.23
DASS Stress	0.15	.03	.27*	0.10-0.21

Note R^2=.59 for Step1; Δ R^2=.66 for Step 2 (ps<.001). *p<.001; DASS: Depression, Anxiety and Stress Scale; EPDS: the Edinburgh Postnatal Depression Scale

Table 2 Factorial loadings for the 10 items of the EPDS generated by explorative factorial analysis – Promin rotation in week 16

EPDS items	Week16	Week25	Week 36	Week9
1. I have been able to laugh and see the funny side of things	.78	.81	.81	.84
2. I have looked forward with enjoyment to things	.76	.81	.84	.83
3. I have blamed myself unnecessarily when things went wrong	.51	.57	.57	.63
4. I have been anxious or worried for no good reason	.70	.73	.73	.75
5. I have felt scared or panicky for no very good reason	.67	.71	.69	.72
6. Things have been getting on top of me	.78	.76	.75	.78
7. I have been so unhappy that I have had difficulty sleeping	.81	.78	.77	.82
8. I have felt sad or miserable	.89	.88	.85	.88
9. I have been so unhappy that I have been crying	.79	.79	.79	.80
10. The thought of harming myself has occurred to me	.81	.73	.67	.77

EPDS: the Edinburgh Postnatal Depression Scale

Table 4. Confirmatory factor analysis for the EPDS, fit indices

	Model tested	Fit indices					
		x^2	df	CFI	TLI	RMSEA (LO90, HI90)	WRMR
1	One-factor model						
	Week 16	554.499	35	.974	.966	.078 (.073-.084)*	2.099
	Week 25	475.909		.973	.965	.079(.072-.085)*	1.941
	Week 36	446.425		.968	.958	.082(.075-.089)*	1.948
2	Model from Jomeen & Martin (2005)						
	Week 16	203.850	19	.987	.981	.063(.056-.071)*	1.491
	Week 25	216.972		.984	.976	.072(.063-.080)*	1.539
	Week 36	239.138		.978	.967	.081(.072-.091)*	1.670
	Model from Jomeen & Martin (2005), bifactor						
	Week 16	190.842	30	.995	.993	.047(.041-.054)*	1.595
	Week 25	181.53		.995	.992	.050(.043-.057)*	1.556
	Week 36	200.788		.992	.989	.057(.050-.065)*	1.636
3	Model from Phillips et al. (2009)						
	Week 16	335.463	34	.985	.980	.061 (.055-.066)*	1.610
	Week 25	315.086		.983	.977	.064(.057-.070)*	1.555
	Week 36	321.709		.978	.970	.069(.063-.076)*	1.638
	Model from Phillips et al. (2009), bifactor						
	Week 16	152.258	30	.997	.995	.041(.035-.048)*	1.425
	Week 25	173.753		.995	.992	.049(.042-.056)*	1.522
	Week 36	178.040		.993	.990	.053(.046-.061)*	1.541
4	Model from Zhong et al., (2014)						
	Week 16	351.825	34	.984	.979	.062(.056-.068)*	1.636
	Week 25	330.722		.982	.976	.066(.059-.072)*	1.584
	Week 36	239.902		.984	.979	.059(.052-.066)*	1.369
	Model from Zhong et al., (2014), bifactor						
	Week 16	89.472	30	.998	.997	.029(.022-.036)*	1.092
	Week 25	103.004		.997	.996	.035(.027-.042)*	1.172
	Week 36	91.424		.997	.996	.034(.026-.042)*	1.104
5	Model from Reichenheim et al. (2011)						
	Week 16	245.396	32	.989	.985	.052(.046-.059)*	1.347
	Week 25	221.314		.988	.984	.054(.047-.061)*	1.281
	Week 36	187.987		.988	.983	.053(.046-.060)*	1.214
	Model from Reichenheim et al. (2011), bifactor						
	Week 16	134.363	26	.997	.995	.041(.035-.049)*	1.338
	Week 25	114.098		.997	.995	.041(.033-.049)*	1.233
	Week 36	96.990		.997	.995	.039(.031-.048)*	1.137

*p<.05
EPDS: the Edinburgh Postnatal Depression Scale
EFA: Explanatory factor analysis
df: degrees of freedom
CFI: Comparative Fit Index
TLI: the Tucker-Lewis index
RMSEA: Root Mean Square Error of Approximation
WRMR: weighted root mean square residual

Table 5. Different cut-off values, their sensitivity, specificity, and positive and negative predictive values of the EPDS.

	Cut	Sens.	Spec.	PV+	PV-	LR+	LR-
EPDS							
Major Depression							
No delay	9	100.00	76.09	31.3	100.0	4.18	0.00
	10	90.00	84.78	39.1	98.7	5.91	0.12
	11	80.00	89.13	44.4	97.6	7.36	0.22
	12	70.00	93.48	53.8	96.6	10.73	0.32
	13	60.00	95.65	60.5	95.6	13.80	0.42
2 to 4 week delay	9	89.23	57.80	19.0	98.0	2.11	0.19
	10	89.23	64.39	21.8	98.2	2.51	0.17
	11	81.54	69.02	22.6	97.1	2.63	0.27
	12	73.85	74.39	24.3	96.2	2.88	0.35
	13	61.54	80.00	25.5	94.9	3.08	0.48

Cut: cut-off values; Sens.: sensitivity; Spec: sprecificity, PV+: positive predictive value; PV-: negative predictive values; LR+: positive likelihood ratio; LR-: negative likelihood ratio; EPDS: the Edinburgh Postnatal Depression Scale

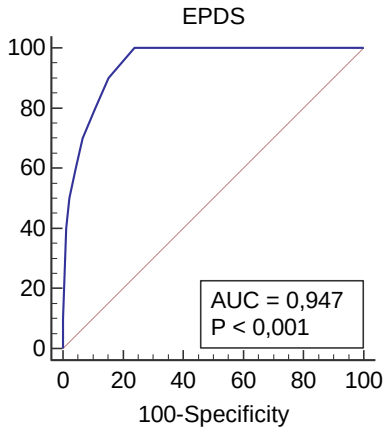

Figure 1 Roc curve for the EPDS in week 16 gestation, no delay

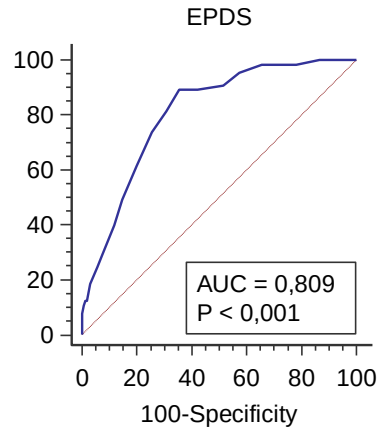

Figure 2 Roc curve for the EPDS in week 16 gestation, 2 – 4 weeks delay